Molecular Biology Reports (2020) 47:2963–2974
https://doi.org/10.1007/s11033-020-05384-9

ORIGINAL ARTICLE

Genome wide association mapping and candidate gene analysis for pod shatter resistance in *Brassica juncea* and its progenitor species

Jasmeet Kaur[1] · Javed Akhatar[1] · Anna Goyal[1] · Navneet Kaur[1] · Snehdeep Kaur[1] · Meenakshi Mittal[1] · Nitin Kumar[1] · Heena Sharma[1] · Shashi Banga[1] · S. S. Banga[1]

Received: 14 January 2020 / Accepted: 19 March 2020 / Published online: 26 March 2020

Abstract

We investigated phenotypic variations for pod shattering, pod length and number of seeds per pod in large germplasm collections of *Brassica juncea* (2n = 36; AABB) and its progenitor species, *B. rapa* (2n = 20; AA) and *B. nigra* (2n = 16; BB). Pod shatter resistance was measured as energy required for rupturing a mature dry pod, with a specially fabricated pendulum machine. Rupture energy (RE) ranged from 3.3 to 11.0 mJ in *B. juncea*. MCP 633, NR 3350 and Albeli required maximum energy to shatter a pod. It ranged from 2.5 to 7.8 mJ for *B. rapa* with an average of 5.5 mJ. *B. nigra* possessed easy to rupture pods. Correlation analysis showed strong associations among these traits in *B. juncea* and *B. rapa.* Genome wide association studies were conducted with select sets of *B. juncea* and *B. rapa* germplasm lines. Significant and annotated associations predict the role of *FRUITFULL, MANNASE7,* and *NAC* secondary wall thickening promoting factor (*NST2*) in the genetic regulation of shatter resistance in *B. juncea*. *NST2* and *SHP1* appeared important for pod length and seeds per pod in *B. rapa*. Candidate gene based association mapping also confirmed the role of *SHP1* and *NST2* in regulating pod shattering and related pod traits in *B. rapa* and *B. juncea*. Footprints of selection were detected in *SHP1*, *SHP2* (*B. rapa*, *B. nigra* and *B. juncea*), *RPL* (*B. rapa*) and *NAC* (*B. juncea*). Our results provide insights into the genetic architecture of three pod traits. The identified genes are relevant to improving and securing crop productivity of mustard crop.

Keywords Pod traits · Candidate gene-based association mapping · Population structure · Loci under selection

Introduction

Pod shattering is an important way of seed dispersal. It is necessary for the reproductive competence of plants under natural conditions but, premature or uncontrolled pod shattering in cultivated crops is undesirable. It restricts the temporal window for crop harvesting and may even lead to crop losses. So, selection for pod indehiscence may have been among the targets of selection during crop domestication. Pod shattering is a complex and intricately controlled process. It involves specific morphological and anatomical features like pod wall valves, a centrally placed replum and valve margins. Pod wall is photosynthetically active fruiting body, which arguably represent a modified leaf [1]. The valve margins include cells which later differentiate to become dehiscence zone (DZ). At maturity, valves separate from replum along the DZ and cause pods to shatter. DZ in shattering resistant species is made up thick walled and lignified parenchyma cells in comparison to thin walled parenchymatous cells present in pod shattering prone species [2]. *Brassicas* crops differ for intensity of pod shattering. *B. juncea*, *B. carinata* and *B. rapa* are less prone to pod shattering as compared to easy to shatter *B. napus*. These variations are attributed to the differences in the pod length [3], size of vascular bundles [4], water content [5], cellulose and lignification in the pod walls [6–8]. Role other anatomical features is also considered important [9, 10]. A suite of regulatory genes, *SHATTERPROOF1* (*SHP1*), *SHATTERPROOF2* (*SHP2*), NAC secondary wall thickening promoting factors (*NST1, NST2* and *NST3*), *INDEHISCENT* (*IND*) and *ALCATRAZ* (*ALC*) are known to regulate pod shattering [11–19]. These are known as valve margin identity genes.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s11033-020-05384-9) contains supplementary material, which is available to authorized users.

✉ S. S. Banga
ssbanga1987@gmail.com

[1] Department of Plant Breeding and Genetics, Punjab Agricultural University, Ludhiana, India

REPLUMLESS (*RPL*) and the *FRUITFULL* (*FUL*) control the expression of valve-margin identity genes [20]. This regulatory network also includes genes associated with leaf development, establishment of dorsoventral axes of the lateral organs (e.g., *FILAMENTOUS FLOWER*, *YABBY3*, *ASYMMETRIC LEAVES1/2*) and the meristematic potential maintenance (*BREVIPEDICELLUS*) [21, 22]. *IND*, *PG* (polygalacturonase) and *FUL* are important for regulating pod shattering resistance in *B. oleracea, B. napus* and *B. juncea* [15, 18, 23]. Multiple QTLs were found to regulate variation for pod shattering in *B. napus* [24] and *B. carinata* [25]. Resistance to pod shattering and the increase in the seed size have been the foci of human selection in *B. juncea* [26].

This article focuses on the variations and genetics of pod shatter resistance in Indian mustard (*Brassica juncea* L. Czern & Coss) and its diploid progenitor species *B. rapa* and *B. nigra*. Indian mustard is a major oilseed and condiment crop in different parts of the world. It is important for India, where the crop is cultivated on over 6.5 million hectares. Majority of the mustard varieties in India are resistant to pod shattering. Although the data for the actual crop losses because of seeds shattered at harvest are not available, yield losses have been reported for many of the newly developed varieties or hybrids in India. Increased propensity for premature shattering of pods in many of these may have resulted from the increased use of exotic germplasm as donors for the novel traits, especially for oil quality. Many exotic germplasm lines and most wild species have easy to shatter pods. Improved resistance to pod shattering is relevant as voluntary seed shattering during traditional hand harvesting leads to crop losses and volunteer pressure during later crop season. Natural variation for shatter resistance was never documented in any global germplasm collection of *B. juncea* or its progenitor species. Information regarding trait genetics is also sketch. In this communication, we report the phenotypic variations for pod shatter resistance in the global germplasm assemblages of *B. juncea* and its progenitor species. We also conducted genome wide and candidate gene-based association mapping to unravel genetic factors underlying the observed trait variation. *SHP1* and *NST2* may have important roles in regulating the variation for all pod traits in the studied crops. We also noted the footprints of selection in *SHP1*, *SHP2* (*B. rapa*, *B. nigra* and *B. juncea*), *RPL* (*B. rapa*) and *NAC* (*B. juncea*).

Materials and methods

Plant materials

Large sets of germplasm lines of *B. juncea* (124) and its progenitor species *B. rapa* (90), *B. nigra* (21) were used for the present studies. These comprised collections from Australia, Canada, China, central Asia, east Europe, Pakistan and India, while *B. rapa* lines were sourced from Canada, Central Asia, Pakistan and India. Details about the germplasm collections are available elsewhere [27]. Experiments were conducted during 2013–2014 (Y1) and 2014–2015 (Y2) cropping seasons at Punjab Agricultural University, Ludhiana. Test genotypes of each species were raised as separate paired rows and replicated twice. Experiments were conducted under timely sown irrigated conditions. Standard agronomic practices were followed throughout the crop season. Germplasm lines under the study were assessed for three pod shattering associated traits. These included: pod strength measured as rupture energy (RE), pod length (PL) and seeds/pod (SPP).

Pod traits

Pod strength was tested with the help of a pendulum machine fitted with an optical encoder [28, 29]. The instrument provides an estimate of RE (mJ) by measuring the loss of movement of the pendulum upon striking and rupturing a pod. The pod is clamped in front of a measuring scale at the bottom dead centre of the pendulum swing. The data were recorded from 25 pods/genotype. The moisture content of all the detached pods was first equilibrated by storing pods at room temperature in plastic tubes containing self-indicating coarse silica gel granules. These were later oven dried at 70 °C for 24 h before measuring their rupture energy. The pod length (mm) was measured on 5 mature pods from 10 random plants and averaged. Seeds from 25 pods from each genotype were counted and averaged to get an assessment of seeds/pod.

Statistical analysis

Mean values of rupture energy, pod length and seeds/pod were calculated for each germplasm and used for further statistical analysis. Analysis of variance (ANOVA) was conducted to study variation due to genotypes, environments (years) and genotype × environment interactions using Minitab Statistical software. ANOVA was conducted independently for each of the three species, e.g. *B. rapa*, *B. nigra* and *B. juncea*.

Genomic DNA extraction and genotyping

Modified CTAB extraction procedure [30] was used for DNA extraction. The quantity and quality of DNA was assessed by agarose gel electrophoreses (0.8%) and nanodrop spectrophotometer. A random set comprising 102 germplasm lines of *B. juncea* were picked for SNP genotyping. *B. juncea* lines were genotyped by sequencing (GBS) [31], while sequencing based DArT was used to

genotype *B. rapa* (78). We also identified six shattering related genes for the candidate gene-based association mapping of the entire germplasm collection (235), comprising *B. nigra*. Gene sequences from *SHP1*, *SHP2*, *NAC*, *IND*, *FUL* and *RPL* were used to develop 21 STS primers. For candidate genes *SHP1* and *SHP2*, specific primers were identified using software PRIMER 3. For *NAC* and *IND*, we used primer sequences as reported by Raman et al. [24] were used. Sequences for *FUL* and *RPL* were retrieved from the *Arabidopsis* database and Blastn with *B. rapa* in the *Brassica* database (BRAD). Orthologus sequences were selected and used for designing primers. For polymorphism studies, test DNA (5 µl of 5 ng/µl) was added to 5.8 µl of master mix that contained 1.0 µl reaction buffer (10 ×), 2.0 µl of 2.0 mM dNTPs, 1.0 µl of 1 mM forward primer, 1.0 µl of 1 mM reverse primer and 0.8 µl Taq polymerase. In vitro amplifications were performed in 96 welled PCR plate in Eppendorf AG (Model 6325). Standard SSR protocol (1 cycle of 4 min at 94 °C; 35 cycles of 1 min at 94 °C, 30 s at T_A, 30 s at 72 °C; 1 cycle of 7 m at 72 °C and a final hold at 4 °C) was followed for PCR analyses. 2.5% Agarose gel (2.5%) was used for electrophoretic separation of the PCR products.

Genome wide association studies

NGSEP-GBS pipeline [32] was used for SNP data calling, based on the reference genome of *B. juncea* v1.5. DArT analytical pipelines were used to handle sequencing data and identification of SNPs in *B. rapa.* Imputation of SNPs was performed with 'Fcgene v1.0.7' [33] and 'Beagle' [34]. Phenotypic data for 102 genotypes of *B. juncea* and 78 genotypes of *B. rapa* was used for GWAS. The data were first normalized by using Johnson transformation available in the software Minitab v16.0. MVP (A Memory-efficient Visualization-enhanced and Parallel-accelerated Tool) [35] and GAPIT [36] were used to measure significance of trait-SNP association, using various GWAS models. An ideal model was expected to show uniformity between the observed and expected p-values in the plot. We matched the observed and expected p values through quantile–quantile (QQ) plots to test the precision of applied GWAS models. PCA + K models were ultimately selected to run with the MLM algorithm as implemented in the MVP software. A threshold of p-value > 3.0 was used for initial detection of significant MTAs. Genomic regions around the identified SNPs were used for annotation. The predicted gene and its orthologous sequences were then annotated by BLAST analysis against the *Arabidopsis thaliana* database using Blast2GO v5.0 tool [37]. Functions of the predicted candidate genes were reviewed in the literature to establish their importance for the traits of interest.

Candidate gene-based association mapping

The full set of 235 germplasm lines {*B. juncea* (124), *B. rapa* (90), *B. nigra* (21)} was used for candidate gene-based association mapping. Population structure study was conducted as implemented in the software STRUCTURE [38]. Candidate gene-based association mapping was carried out with the software package TASSEL (Trait Analysis by Association, Evolution and Linkage) [39].

Population genetic analysis

These estimations were based on polymorphisms generated by the primers designed from sequence information of six pod shattering related candidate genes. Expected heterozygosity and deviations from the Hardy–Weinberg equilibrium were worked out with the help of software FSTAT 2.9.3.2 [40]. Other population genetic parameters were measured with the help of the software package ARLEQUIN [41–43].

Results

Analysis of variance (ANOVA) showed significant genotypic differences for three pod traits all three species. Variances due to environment were significant for RE and seeds/pod in *B. rapa* and *B. juncea* and pod length in *B. juncea*. Genotypes × environment interactions (G × E) were significant for RE in *B. rapa* and *B. juncea;* seeds/pod and pod length for the test species (Table 1). Environment effects were non-significant for rupture energy in *B. nigra*, but G × E interactions variances were significant for pod length and seeds/pod. Estimates for pod strength, pod length and seeds/pod are presented in Table 2.

Pod shattering (rupture energy)

RE values in *B. rapa* ranged from 2.3 (Leega) to 9.0 mJ (Mitra) during Y1 and 2.7 (Leega) to 7.7 mJ (Texi) during Y2. Variation was high for rupture energy in the *B. rapa* collection from central Asia, it was low for brown sarson germplasm from India (Fig. S1). *B. nigra* appeared very prone to shattering with little genotypic variation. RE values ranged from 1.3 mJ (Assam) to 2.6 mJ (Gujrat) during Y1, average being 1.9 mJ. Y2 values ranged from 1.4 mJ (Assam, FRG1) to 2.4 mJ (R 4136) with an average of 1.9 mJ (Fig. S2). *B. juncea* genotypes showed the largest estimates for RE; these varied from 2.9 mJ (EC 564645) to 11.5 mJ (MCP 633), with an average of 6.5 mJ during Y1, and from 3.4 mJ (RM 51) to 10.7 mJ (MCP 633), with an average of 6.3 mJ, during Y2. Australian *B. juncea* germplasm lines were more sensitive to

Table 1 Analysis of variance (ANOVA) for different pod related traits in evaluated three *Brassica* species

Source	Rupture energy			Pod length			Seeds/pod		
	B. rapa	*B. nigra*	*B. juncea*	*B. rapa*	*B. nigra*	*B. juncea*	*B. rapa*	*B. nigra*	*B. juncea*
Replication	0.14	0.022	0.257	0.74	0.008	16.34	0.019	0.002	2.942
Environment (years)	15.76**	0.042	10.54**	0.002	1.0296	138.14**	172.14**	0.349	59.80**
Genotype	5.26**	0.36**	10.44**	72.94**	3.10**	86.34**	36.51**	3.40**	15.81**
Genotype x Environment	0.41**	0.013	0.53**	23.64**	2.70**	9.37**	16.94**	1.86*	5.09**
Error	0.19	0.033	0.286	9.3	0.566	5.5	3.29	0.816	1.311
DF	89	20	124	89	20	124	89	20	124

*Significant at 0.05 level

**Significant at 0.01 level

shattering and also had a narrow range of variation. Distribution of variation in Indian and east European genotypes was skewed towards high RE values (Fig. S3).

Pod length

Maximum variation for PL was recorded in *B. rapa*. These varied from 17.0 mm (PAK 85530) to 48.0 mm (Sanya) for Y1 and 24.0 mm (R 436) to 42.8 mm (PAK 85910) during Y2. PL average was 34.2 mm. Based on geographic groupings, maximum trait variation was recorded from Pakistan, but a narrow range of variation was observed for Indian brown sarson (Fig. S1). PL was much smaller in *B. nigra* as compared to *B. rapa* and *B. juncea.* These fell within a range of 9.0 mm (R 704) to 15.0 mm (CN 113784) during Y1 because these varied between 10.0 mm (R 4130) to 13.0 mm (R 4132) for Y2. The corresponding averages were 11.7 mm and 11.3 mm. Canadian germplasm lines showed a wider range for the pod length, while Indian *B. nigra* lines had a particularly narrow range (Fig. S2). Pod length ranged from 18.5 (AJ 1) mm to 44.0 mm (Rohini) in *B. juncea* with an average of 30.6 mm during Y1. This was compared to 21.0 mm (EC 564,649) to 47.8 mm (MCP 633) recorded for Y2, the average being 31.6 mm. At the population level, Australian genotypes varied the least (Fig. S3). Variation in the Indian germplasm lines was skewed towards longer pods.

Seeds/pod

Brassica rapa showed a wide range for the seeds/pod trait. During Y1, the seeds/pod ranged from 5.8 (R 436) to 28.5 (Qianxi Hei) with an average of 15.2 seeds/pod. The range for Y2 was narrower, 5.8 (R 436) to 23.6 (German 56), average being 13.9 seeds/pod. Maximum variation was recorded in the germplasm lines from Central Asia while a narrow range of variation range was recorded in Indian brown sarson (Fig. S1). *B. nigra* showed low variation, with values of 4.0 (UP) to 8.0 (R 4136) seeds/pod during Y1 and 4.0 (UP) to 7.5 (R 4134) for Y2 (Fig. S2). The number of seeds/pod in *B. juncea* ranged from 7.0 (AJ 2) to 20.1 (IC 597895) seeds/pod with an average of 13.8 seeds/pod, during Y1. For Y2, these values ranged between 8.6 (RE 8) to 18.5 (IC 597875) seeds/pod. Australian *B. juncea* had least variation (Fig. S3). Pod shatter energy was associated with pod length in both *B. rapa* (0.251) and *B. juncea* (0.489). It was negatively correlated with seed number (− 0.273) and pod length (− 0.176) in *B. nigra*.

Association mapping for pod shattering and related traits

GWAS was conducted to find marker trait associations (MTAs) to analyze genes linked to the rupture energy, pod length and seeds/pod (Table 3) (Figs. 1, 2). Twenty-three MTAs associated with RE were envisaged on chromosomes A02, A03, A05, A09 and B05 of *B. juncea.* Phenotypic variation explained ranged between 20.67 and 25.46%. Three SNPs A02_15096881, A02_15096938, A02_15097017 was present near NAC domain transcriptional regulator superfamily protein, *AT1G76420-CUC3/NAC368* (*CUP SHAPED COTYLEDON 3*). We could also observe the *AT3G61910-NST2* along the SNPs, A09_46982925 and A09_46982955. A03_13393564 was associated with *AT2G45190-AFO* (*ABNORMAL FLORAL ORGANS*)/*FIL* (*FILAMENTOUS FLOWER*). *AT5G60910-FUL1/AGL8* (*FRUITFULL 1*/*AGAMOUS-LIKE 8*) was called close to SNP A03_27065868. Two copies each of ethylene responsive genes, *ERF4* (*ETHYLENE RESPONSIVE ELEMENT BINDING FACTOR 4*) and *RAP2.4* (*ETHYLENE-RESPONSIVE TRANSCRIPTION FACTOR RAP2-4*) were seen close to SNPs A05_12315809, A05_12315841 and B05_60868573, B05_60868600, B05_60868609, B05_60868738, B05_60868793, B05_60868821, B05_60868825, respectively. Another important gene *AT5G66460-MAN7* (*ENDO-BETA-MANNASE7*) was along the A09_6549794. It encodes glycosyl hydrolase superfamily protein that serves in the development and dehiscence of *Arabidopsis* pods. We anticipated no gene for RE in *B. rapa.* We could predict the importance of

Table 2 Mean values for rupture energy, pod length and seeds/pod in *B. rapa*, *B. nigra* and *B. juncea*

Species	Rupture energy (mJ)			Pod length (mm)			Seeds/pod		
	Year 1	Year 2	Pooled	Year 1	Year 2	Pooled	Year 1	Year 2	Pooled
	Mean ± SE (range)								
B. rapa									
rap CA	4.9 ± 0.3 (2.3–7.8)	4.9 ± 0.2 (2.7–7.2)	4.9 ± 0.3 (2.5–7.3)	32.3 ± 0.6 (19.0–44.5)	33.2 ± 0.4 (25.3–37.8)	32.7 ± 0.5 (25.5–40.0)	14.7 ± 0.5 (7.0–28.5)	14.5 ± 0.5 (6.1–23.6)	14.6 ± 0.5 (8.3–24.1)
rap CAN	5.8 ± 0.3 (4.3–8.0)	5.4 ± 0.2 (4.3–7.7)	5.6 ± 0.3 (4.3–7.8)	34.3 ± 1.6 (24.0–48.0)	33.6 ± 0.9 (24.0–41.3)	34.0 ± 1.1 (24.0–42.9)	15.5 ± 1.4 (5.8–24.5)	12.7 ± 0.7 (5.8–17.3)	14.1 ± 0.9 (5.8–20.1)
rap IT	5.9 ± 0.5 (3.1–9.0)	5.3 ± 0.2 (3.5–6.2)	5.6 ± 0.4 (3.3–7.8)	34.0 ± 1.4 (25.5–40.0)	35.1 ± 0.9 (29.0–38.3)	34.6 ± 1.1 (27.3–38.9)	13.6 ± 0.7 (10.5–17.9)	13.9 ± 0.5 (11.2–16.5)	13.7 ± 0.5 (11.0–16.1)
rap PB	6.1 ± 0.2 (2.9–8.0)	5.6 ± 0.1 (3.0–7.3)	5.8 ± 0.2 (3.0–7.6)	35.2 ± 0.9 (17.0–46.5)	34.9 ± 0.6 (25.3–42.8)	35.0 ± 0.7 (21.1–44.6)	15.7 ± 0.6 (9.7–19.2)	14.1 ± 0.4 (9.7–19.2)	14.9 ± 0.4 (9.3–18.9)
rap IB	4.8 ± 0.3 (4.4–5.5)	4.6 ± 0.4 (3.9–5.4)	4.7 ± 0.4 (4.1–5.5)	34.3 ± 0.6 (34.0–35.5)	31.9 ± 0.9 (30.3–33.3)	33.1 ± 0.3 (32.8–33.6)	17.4 ± 1.0 (16.3–19.5)	13.2 ± 0.4 (12.4–13.9)	15.4 ± 0.7 (14.5–16.7)
Overall mean	5.7 ± 0.1	5.3 ± 0.1	5.5 ± 0.1	34.2 ± 0.6	34.2 ± 0.4	34.2 ± 0.4	15.2 ± 0.4	13.9 ± 0.3	14.5 ± 0.3
B. nigra									
nig CAN	1.9 ± 0.1 (1.4–2.4)	1.9 ± 0.1 (1.4–2.4)	1.9 ± 0.1 (1.4–2.3)	11.6 ± 0.4 (9.0–15.0)	11.4 ± 0.2 (10.0–13.0)	11.4 ± 0.2 (10.0–13.0)	6.5 ± 0.3 (4.4–8.0)	6.3 ± 0.2 (5.2–7.5)	6.4 ± 0.1 (5.0–7.4)
nig I	2.0 ± 0.3 (1.3–2.6)	1.9 ± 0.3 (1.4–2.2)	2.0 ± 0.3 (1.4–2.4)	12.0 ± 0.6 (11.0–13.0)	10.7 ± 0.3 (10.0–11.0)	11.3 ± 0.3 (11.0–12.0)	4.7 ± 0.4 (4.0–5.5)	4.7 ± 0.5 (4.0–5.7)	4.7 ± 0.4 (4.0–5.2)
Overall mean	1.9 ± 0.1	1.9 ± 0.1	1.9 ± 0.1	11.7 ± 0.3	11.3 ± 0.2	11.4 ± 0.2	6.3 ± 0.3	6.1 ± 0.2	6.2 ± 0.2
B. juncea									
jun A	4.9 ± 0.6 (4.0–7.8)	4.7 ± 0.4 (4.1–6.9)	4.8 ± 0.5 (4.1–7.3)	27.1 ± 0.9 (24.0–30.5)	28.1 ± 1.1 (24.3–32.3)	27.6 ± 0.6 (26.1–29.5)	13.9 ± 0.2 (13.5–14.5)	13.9 ± 1.0 (11.7–18.4)	13.9 ± 0.5 (12.6–15.9)
jun C	5.4 ± 0.5 (4.4–8.1)	4.9 ± 0.4 (4.1–7.0)	5.2 ± 0.4 (4.3–7.5)	22.2 ± 1.5 (18.5–30.0)	24.7 ± 0.9 (22.5–30.0)	23.4 ± 1.2 (21.1–30.0)	10.7 ± 1.1 (7.0–14.1)	11.3 ± 0.8 (9.0–15.2)	11.0 ± 0.9 (8.5–14.6)
jun EE	6.0 ± 0.2 (2.9–11.1)	5.8 ± 0.2 (3.5–10.0)	5.9 ± 0.2 (3.4–10.5)	29.3 ± 0.5 (20.5–41.0)	30.1 ± 0.5 (21.0–38.3)	29.7 ± 0.5 (21.0–39.6)	14.2 ± 0.4 (8.9–20.1)	13.1 ± 0.3 (8.6–18.5)	13.7 ± 0.3 (8.7–17.8)
jun I	7.4 ± 0.2 (3.2–11.5)	7.1 ± 0.2 (3.4–10.7)	7.2 ± 0.2 (3.3–11.0)	33.3 ± 0.6 (24.0–44.0)	34.5 ± 0.6 (25.0–47.8)	33.9 ± 0.5 (27.3–44.9)	13.7 ± 0.2 (7.1–16.8)	13.1 ± 0.2 (8.9–18.0)	13.4 ± 0.2 (9.8–15.9)
Overall mean	6.5 ± 0.2	6.3 ± 0.1	6.4 ± 0.1	30.6 ± 0.4	31.6 ± 0.4	31.1 ± 0.4	13.8 ± 0.2	13.1 ± 0.2	13.4 ± 0.2

AT2G46770-NST1 in describing variation for pod length in *B.juncea*. The gene was envisaged near seven closely placed SNPs in the genomic region (11,913,579–11,913,680) on the chromosome B06. We also envisaged two important genes, *AT3G58780-SHP1* (*SHATTERPROOF1*) and *AT3G61910-NST2* close to the SNP A09_45505210 in *B. rapa*. Both these genes were also showed for seed per pod, where these genes were predicted close to A09_45507062. For seeds/pod in *B.juncea, AT2G20610-SUR1* (*SUPERROOT1*) was envisaged near a cluster of nine SNPs in the genomic region 4,867,097–4,867,166 on the chromosome B02. It controls auxin biosynthesis. Candidate gene based association mapping did not lead to identification of any significant marker-trait associations in *B. rapa* and *B. nigra* for rupture energy (Table 4). However, marker *NAC*_3 explained 20–26% of variation for pod length in *B. rapa*. *NAC*_1, *SHP* 1-2_1 and *SHP* 1-3_1 in *B. rapa* and *SHP* 1-7_1 in *B. nigra* showed association with the seeds/pod. In *B. juncea*, *SHP* 1-3_1 and *SHP* 2-2_1 were associated with rupture energy, whereas markers *SHP* 1-3_1 and *SHP* 2-9_1 were associated with pod length. *SHP* 1-9_1 and *NAC*_1 were associated with the seeds/pod.

Table 3 GWAS for traits associated with pod shatter resistance

Species	Trait	Chr	SNP	BP	P.V.	Effect	Consensus	Gene	Function
B. juncea	RE	2	A02_15096881, A02_15096938, A02_15097017	15,096,881–15,097,017	21.48	0.65		*AT1G76420-CUC3/NAC368 (CUP SHAPED COTYLEDON3)*	Shoot meristem and pod formation
		3	A03_13393564	13,393,564	21.94	2.53	PL	*AT2G45190-AFO/FIL (ABNORMAL FLORAL ORGANS)*	Shoot apical meristem organization and alters the activity of *AP2*
		3	A03_27065868	27,065,868	25.46	− 1.96		*AT5G60910-FUL1 (FRUITFULL)*	Expression of the valve margin identity genes (*SHP1* and *SHP2*)
		5	A05_12315809, A05_12315841	12,315,809–12,315,841	21.57	− 0.97		*AT3G15210-ERF4 ((Ethylene-responsive transcription factor4)*	Contain one *APETALA2* domain which plays role in ethylene signaling
		9	A09_6549794	6,549,794	21.91	− 2.24		*AT5G66460-MAN 7 (ENDO-BETA-MANNASE7)*	Dehiscence of pods
		9	A09_13128924, A09_13128984, A09_13128997, A09_13129099	13,128,924–13,129,099	21.00	0.83		*AT1G62360-STM (SHOOT MERISTEMLESS)*	Shoot apical meristem (SAM) function
		9	A09_13105514	13,105,514	20.67	− 1.80		*AT1G62360-STM (SHOOT MERISTEMLESS)*	Shoot apical meristem (SAM) function
		9	A09_46982925, A09_46982955	46,982,925–46,982,955	21.20	− 2.18		*AT3G61910-NST-2 (0.3 Mb)*	Formation of the pod
		15	B05_60868573, B05_60868600, B05_60868609, B05_60868738, B05_60868793, B05_60868821, B05_60868825	60,868,573–60,868,825	20.83	1.25		*AT1G78080-RAP2.4 (Ethylene-responsive transcription factor RAP2.4)*	Contain one *APETALA2* domain which plays role in ethylene signaling
		15	B05_4478265	4,478,265	21.26	− 1.74		*AT4G32551 LEUNIG*	Alter *AP2* expression
	PL	16	B06_11913579, B06_11913585, B06_11913594, B06_11913604, B06_11913643, B06_11913670, B06_11913680	11,913,579–11,913,680	30.55	− 0.82		*AT2G46770 NST1 (0.5 Mb away) (NAC family protein)*	Formation of pod
	SPP	12	B02_4867097, B02_4867098, B02_4867115, B02_4867125, B02_4867126, B02_4867131, B02_4867143, B02_4867160, B02_4867166	4,867,097–4,867,166	21.06	0.11		*AT2G20610-SUR1 (SUPERROOT1)*	Pod dehiscence
B. rapa	PL	9	A09_45505210	45,505,210	18.84	− 2.07	SPP	*AT3G58780-SHP1 (SHATTERPROOF1) (0.5 Mb), AT3G61910-NST2 (0.8 Mb)*	Formation of the pod; pod dehiscence
	SPP	9	A09_45507062	45,507,062	16.71	− 1.65		*AT3G58780-SHP1 (SHATTERPROOF1) (0.5 Mb), AT3G61910-NST2 (0.8 Mb)*	Formation of the pod; pod dehiscence

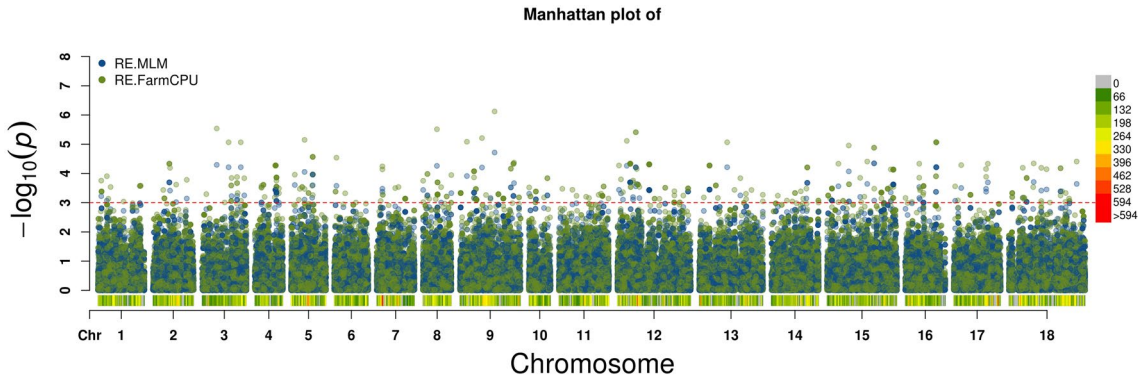

Fig. 1 Manhattan plots depicting the MTAs for rupture energy in *B. juncea*

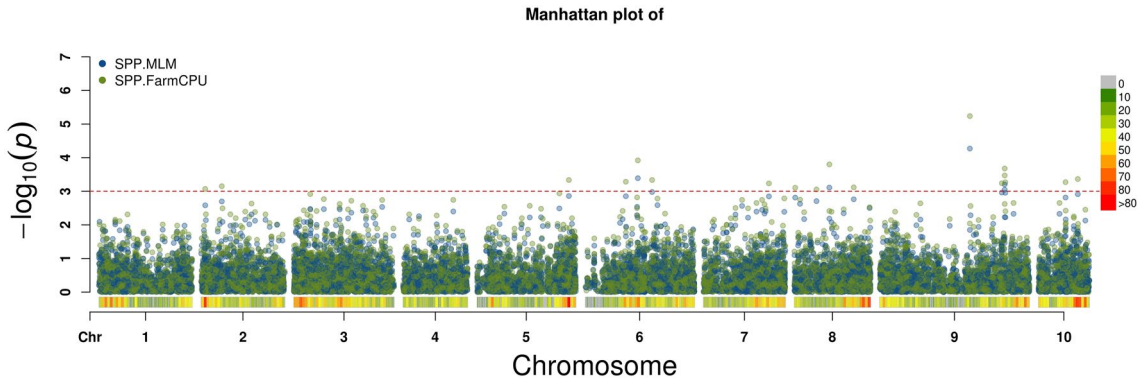

Fig. 2 Manhattan plots depicting the MTAs for seeds per pod in *B. rapa*

Table 4 Marker-trait association in *B. rapa*, *B. nigra* and *B. juncea*

Species	Marker loci	Trait	Year 1		Year 2		Pooled	
			− log10 (P)	R^2	− log10 (P)	R^2	− log10 (P)	R^2
B. rapa	NAC_3	Pod length	3.71	26.41	–	–	2.65	20.45
	NAC_1	Seeds/pod	–	–	3.43	13.26	2.72	11.19
	SHP 1-3_1	Seeds/pod	–	–	2.66	15.75	2.01	12.35
	SHP 1-2_1	Seeds/pod	–	–	2.41	16.41	2.35	15.48
B. nigra	SHP 1-7_1	Seeds/pod	–	–	2.36	20.64	2.19	18.45
B. juncea	SHP 1-3_1	Rupture energy	–	–	–	–	2.19	8.16
	SHP 2-2_1	Rupture energy	–	–	–	–	1.96	7.24
	SHP 1-3_1	Pod length	2.41	7.24	2.8	7.25	2.72	8.19
	SHP 2-9_1	Pod length	2.27	6.18	2.3	6.37	2.6	7.45
	SHP 1-9_1	Seeds/pod	3	18.44	2.54	9.76	3.1	11.21
	NAC_1	Seeds/pod	–	–	3	18.22	2.56	16.55

Genetic diversity, population structure and linkage disequilibrium

Allelic variation for six pod shattering associated candidate genes was investigated. Observed heterozygosity (H_O) was low in *B. rapa* (Table 5). In contrast, observed heterozygosity was higher than the expected heterozygosity (H_E) in *B. nigra*. Number of alleles per locus (N_A) ranged from 1.3 to 2.1 in *B. rapa*, maximum being in rapa PB. Lowest value was recorded for rapa IB. The mean effective estimate of alleles per locus (N_E) ranged from 1.2 to 1.6, rapa CA had the maximum value. Canadian genotypes of *B. nigra* recorded maximum values for N_A and N_E. Differences in mean values of N_A, N_E, H_O, H_E and μH_E (unbiased expected heterozygosity)

Table 5 Genetic diversity estimates (mean + SE) at the population levels based on candidate gene loci

Species	Pop	N_A	N_E	H_O	H_e	μH_e
B. rapa	rapa CA	1.9	1.6	0.32	0.32	0.32
	rapa CAN	2.0	1.5	0.26	0.28	0.29
	rapa IT	1.8	1.5	0.22	0.26	0.28
	rapa PB	2.1	1.5	0.22	0.27	0.28
	rapa IB	1.3	1.2	0.17	0.15	0.19
	Average	1.8	1.5	0.24	0.26	0.27
B. nigra	nig CAN	2.0	1.8	0.50	0.40	0.42
	nig I	1.5	1.4	0.39	0.26	0.32
	Average	1.8	1.6	0.45	0.33	0.37
B. juncea	jun A	1.7	1.6	0.54	0.33	0.36
	jun C	1.8	1.7	0.56	0.37	0.41
	jun EE	1.9	1.6	0.47	0.35	0.35
	jun I	1.9	1.7	0.45	0.36	0.36
	Average	1.8	1.7	0.51	0.35	0.37

N_A number of alleles per locus, N_E mean effective number of alleles per locus, H_O observed heterozygosity, H_e expected heterozygosity, μH_e unbiased expected heterozygosity

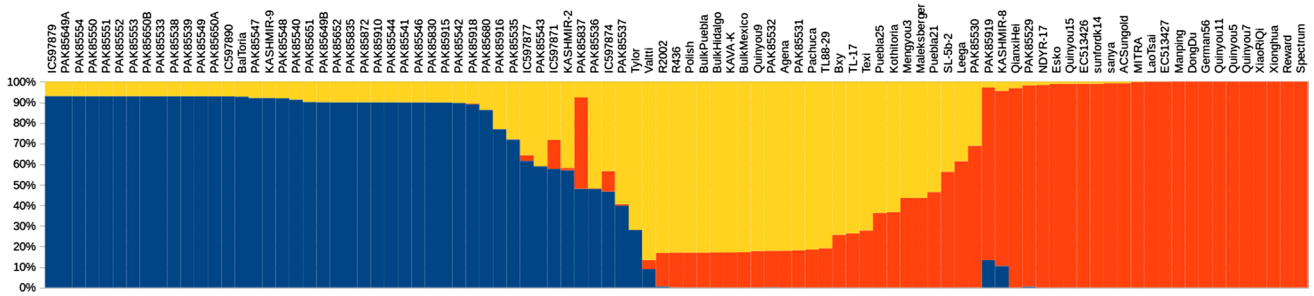

Fig. 3 Population structure showing three populations of *B. rapa* based on software STRUCTURE

were not significant ($P \leq 0.05$) among studied populations. Structure analysis helped to group *B. rapa* germplasm into three subpopulations at $\Delta K = 3$ (Fig. 3). Significant admixing was showed; exception was the group (portrayed in red colour) comprising germplasm from central Asia. Indian brown sarson, Indian toria and Canadian *rapa* were present as admixtures. Group with blue colour was made up primarily of brown sarson germplasm from Pakistan and few brown sarson and toria lines of Indian origin. Group with yellow colour had genotypes from Canada, Central Asia and few toria genotypes from Pakistan and India. The distribution for all the three traits was normal in structure defined groups (Fig. S4). The genotypes included in the blue group varied in a narrow range for rupture energy and pod length as compared to other groups. Population structure was absent in *B. nigra* with complete admixing (Fig. 4). *B. juncea* accessions formed five subpopulations at $\Delta K = 5$ (Fig. 5). East European lines predominated the first group (brown colour). Blue group included eleven lines from east Europe, seven lines from India and two lines each from Canada and Australia.

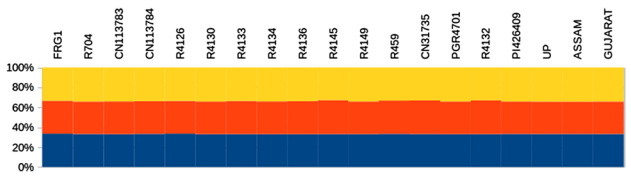

Fig. 4 Population structure of *B. nigra* based on software STRUCTURE

Group III (green colour) included 10, 13 and 1 accession(s) from east Europe, India and Canada. Group with red colour included three Canadian, four Australian, eleven east European and five Indian germplasm lines. Indian accessions (27 out of 51) predominated the yellow group. Despite few exceptions, the distribution spread for rupture energy in structure defined groups showed a normal and overlapping distribution (Fig. S4). For pod length, four out of five structure defined populations had a normal distribution. Bulk of Indian genotypes included in a yellow group showed a

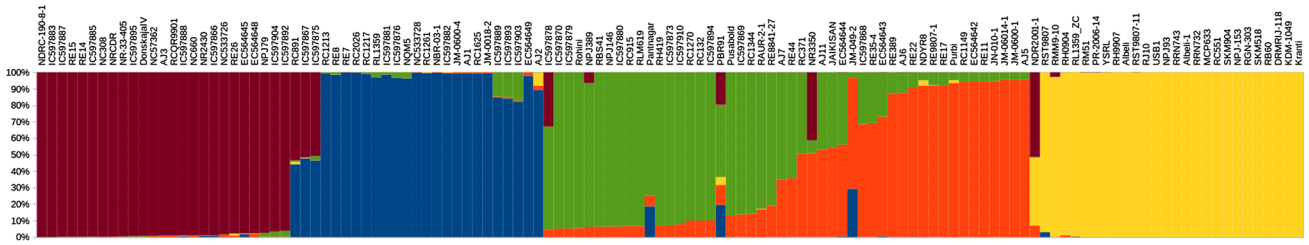

Fig. 5 Population structure showing five subpopulations of *B. juncea* based on software STRUCTURE

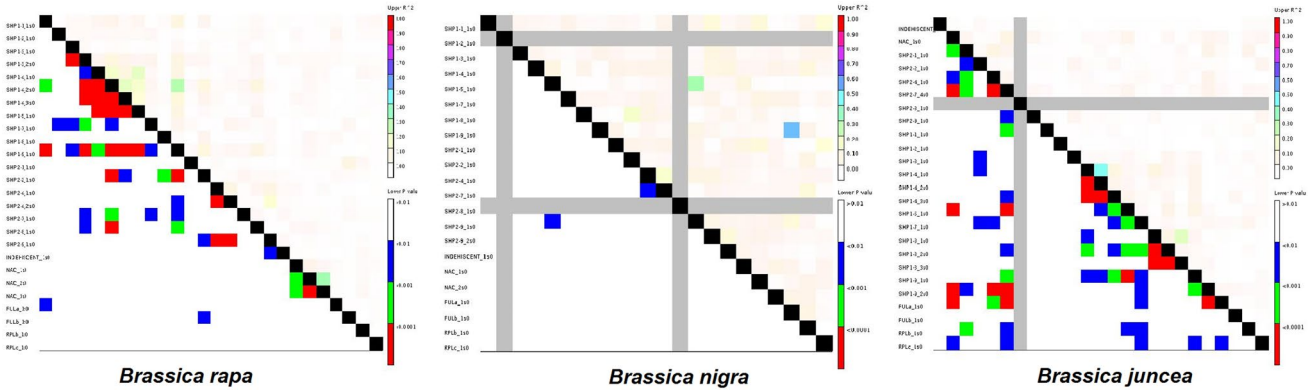

Fig. 6 Decay of linkage disequilibrium (LD) in *B. rapa, B. nigra* and *B. juncea*

distribution pattern that was skewed towards higher values (Fig. S4). Only green group had distribution that was skewed towards a higher number of seeds/pod as compared to other groups, where the variation pattern was largely normal (Fig. S4). The average LD varied between species and subpopulations within the species (Fig. 6). LD blocks differed across *B. rapa* subpopulations. Two significant LD blocks were recorded in Indian brown sarson, spanning *SHP1* and *FUL*. Weak LD blocks, involving *SHP1* and *NAC*, were observed in Canadian *rapa* and Indian toria. LD block was also seen for *SHP1* in Pakistani brown sarson. *NAC* figured in a small LD block (Fig. S5). LD blocks also differed in *B. nigra*. Canadian *B. nigra* had a weak LD blocks involving *SHP2*, whereas LD block in Indian types spanned both *SHP1* and *SHP2* (Fig. S6). Weak LD blocks were noticed for *SHP1* in Australian and Chinese *B. juncea*. Indian *B. juncea* germplasm showed a big LD block for *SHP1* and a smaller one for *SHP2*. East European *B. juncea* also showed significant LD blocks for *SHP1* (Fig. S7).

Detection of loci under selection

We compared values for fixation indices (F_{ST}) with observed heterozygosity to show loci under selection (Table 6, Fig. S8). Loci with low F_{ST} values were under selection for less than the expected heterozygosity. Different loci under selection were identified in three *Brassica* species. These were: *SHP1, SHP2*, *RPL* in *B. rapa; SHP1*, *SHP2* in *B. nigra* and *SHP1*, *SHP2* and *NAC* in *B. juncea*.

Table 6 Analysis of molecular variance

Dataset	Source of variation	Percentage variation	F_{st}
B. rapa	Among populations	15.17	0.152
	Within populations	84.83	
B. nigra	Among populations	6.69	0.067
	Within populations	93.31	
B. juncea	Among populations	4.36	0.044
	Within populations	95.64	

F_{ST} is genetic differentiation based on allele identity

Discussion

QTL mapping and GWAS have been widely used to identify genes controlling pod dehiscence in many crop plants [1, 24, 44]. However, little is known about the variation and genetics of this trait in *B. juncea*. We used multiple approaches to investigate variation, population structure and effect of selection on genetic diversity for pod traits in *B. juncea,*

relative to its progenitor species. We recorded a wide range and near normal distribution of variation for RE, pod length and seeds/pod in *B. juncea* and *B. rapa*. *B. nigra* varied the least. Distribution of variations was almost similar in *B. juncea* and *B. rapa* germplasms. This may indicate that these two species have adapted to similar selection pressures for pod traits. The estimates of observed heterozygosity were lower than expected heterozygosity in *B. juncea* and *B. rapa*. Reverse was true for *B. nigra*, where observed heterozygosity was higher than expected heterozygosity. This may be a consequence of admixing of isolated populations or lack of selection pressure for the trait in this species. Prevention or reduction of natural seed dispersal was possibly among the first traits targeted for selection by ancient domesticators. This coupled with strong selection pressure exerted by modern plant breeders may have caused partial or quantitative loss of seed dispersal and shifts in gene frequencies. We could identify many germplasm lines of *B. rapa* and *B. juncea* with very hard to shatter pods. Low shattering *B. rapa* germplasm may be useful as trait donor for improving pod shatter resistance in *B. napus*. Indian and east European accessions of *B. juncea* were less prone to pod shattering as compared to Australian genotypes. Pod rupture energy was correlated with pod length in *B. rapa* and *B. juncea*, also reported earlier in *B. napus* [45].

Candidate genes associated with pod traits

GWAS led to the identification of 23 MTAs for pod shatter resistance in *B. juncea*, with high phenotypic variation explained. MTAs were identified on different chromosomes of both A (A02, A03, A05, A09) and B (B02, B05 and B06) genomes. Annotations of the nearby genomic regions allowed us to identify important pod development and related genes. *NAC* domain transcriptional regulator superfamily protein, *AT1G76420-CUC3/NAC3* was identified on chromosome A02. Plant-specific transcription factors of *NAC* family includes, secondary wall thickening promoting factors 1, 2 and 3 (*NST1*, *NST2* and *NST 3*) [46]. These regulate development of the pods in *Arabidopsis*. *AT3G61910-NST2* was recognized on chromosome A09 and *NST1* was annotated close to the SNPs identified as significant for pod length on chromosome B06. *NST1* regulates the development of pods in *Arabidopsis*. *AT2G45190-AFO/FIL*, plant-specific transcription factor of YABBY family protein, was predicted on A03. Two YABBY family transcription factors *FILAMENTOUS FLOWER* (*FIL*) and *YABBY3* (*YAB3*) are critical for the valve margin development and they act in tandem to promote expression of *FUL* in the valves, and *SHP1,2* in the valve margins [47]. *FIL* also controls the spatial activity of *APETALA2*-like transcription factor which impacts the ease of threshing [48]. We envisaged *AT5G60910-FUL1/AGL8* on chromosome A03. This gene encodes a MADS-box transcription factor [49] that mediates valve development by inhibiting expression of the valve margin identity genes (*SHP1* and *SHP2*) [19, 20]. *FUL* is required for pod morphogenesis, and its expression is integral to *SHP* genes [49, 50]. Studies have shown that the ectopic expression of the *FUL* may lead to shatter proof fruit in *B. juncea* [13]. Ethylene responsive genes, *ERF4* and *RAP2.4* were envisaged on chromosomes A05 and B05. Both these involve one *APETALA2* domain which mediates light and ethylene signaling [51]. Up-regulation of ABA signaling and down-regulation of ethylene and jasmonate signaling also reduced pod shattering [6, 52]. *APETALA2* has been reported to specify abscission zone (AZ) development and seed shattering in rice [53]. *AT5G66460-MAN7* was envisaged on the chromosome A09. It is a glycosyl hydrolase superfamily protein that functions in the development and dehiscence of *Arabidopsis* pods. Its expression in the pod is depended on the *IND* and *ALC* transcription factors. *MAN7* is also perceived to promote cell degeneration in the separation layer in mature pods [54]. We also envisaged two important genes, *AT3G58780-SHP1* and *AT3G61910-NST2* on chromosome A09 of *B. rapa*. *NST2* regulates the pod formation. *AT3G58780-SHP1* along with *SHP2* controls DZ differentiation and also promotes the lignification of adjacent cells [14]. These are fundamental for the coordination of cell divisions in ovule, seed coat development and endosperm formation in *Arabidopsis* [55]. *AT2G20610-SUR1*, controlling the auxin biosynthesis, was predicted close to a cluster of SNPs on B02. Low auxin level is essential for pod dehiscence as it may trigger cell wall degrading enzymes [17]. Our candidate gene-based association studies validated some of the gene predictions from GWAS. Both *SHP1* and *SHP2* were associated with rupture energy in *B. juncea*. *SHP1* (all three species) and *NAC1* (*B. rapa* and *B. juncea*) were associated with seeds/pod. We also predicted more than one copy of shattering related gene(s), either on the same chromosome or on different chromosomes. Multiple copies of shatter related genes are expected as *Brassicas* are ancient polyploids. Further, an expansion in gene families was also possible because of the direct or indirect selection pressure.

Linkage disequilibrium and footprints of selection

Analysis of allelic diversity in the candidate genes emphasized existence of population structure in *B. rapa* and *B. juncea*. Gene exchange/admixing, observed in *B. rapa*, was expected in a cross-pollinated crop. We also identified LD blocks in *B. rapa*. However, these differed across subpopulations of the species. These spanned *SHP1* and *FUL* in brown sarson, while Canadian rapa and Indian toria showed weak LD blocks involving *SHP1* and *NAC*. In contrast, *B. nigra* showed weak LD blocks involving *SHP2* or both *SHP1&2*. Indian and east European *B. juncea* formed large

LD blocks for *SHP1*. In contrast, Australian and Chinese *B. juncea* revealed weak LD blocks for the gene. Australian and Chinese genotypes were more susceptible to pod shattering as compared to Indian and east European germplasm. It is likely that diverse sets of genes were selected in different selection events and ecogeographic groups. LD is also influenced by extent of genetic diversity present in a germplasm resource. High genetic variation is associated with a rapid LD decay; a direct consequence of the wide-ranging historical recombination's. Random genetic drift in finite populations can also generate regions of high LD [56]. Intensive selection reduces natural variation and heterozygosity inherent in the species. *SHP1*, *SHP2* (*B. rapa*, *B. nigra* and *B. juncea*), *RPL* (*B. rapa*) and *NAC* (*B. juncea*) carried footprints of selection. These genes were also predicted from our GWAS studies for their role in regulating rupture energy, pod length and seeds/pod.

Conclusions

SHP1 and *NAC* appears useful for developing breeder friendly KASP (Kompetitive Allele-Specific PCR) assays for marker aided improvement of pod traits in mustard. Mustard genotypes MCP 633, NR 3350 and SKM518 can be used as trait donors for pod length, seed number and resistance to pod shattering.

Acknowledgements We thank Indian Council of Agricultural Research for financial support of this study that was conducted within under ICAR National Professor Project "Broadening the genetic base of Indian mustard (*Brassica juncea*) through alien introgressions and germplasm enhancement" awarded to S.S.B. The authors also thank "Plant Gene Resources of Canada, Agriculture and Agri-Food Canada" for providing much of the germplasm used in these studies.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Ogutcen E, Pandey A, Khan MK, Marques E, Varma-Penmetsa R, Kahraman A, von Wettberg EJB (2018) Pod shattering: a homologous series of variation underlying domestication and an avenue for crop improvement. Agronomy 8:137
2. Kaur N, Setia N, Setia RC, Banga SS (2009) Comparative studies of morphological and anatomical characteristics of siliqua in shatter susceptible and resistant *Brassica* cultivars. Phytomorphology 59:116–120
3. Summers JE, Bruce DM, Vancanneyt G, Redig P, Werner CP, Morgan C, Child RD (2003) Pod shatter resistance in the resynthesized *Brassica napus* line DK142. J Agric Sci 140:43–52
4. Child RD, Summers JE, Babij J, Farrent JW, Bruce DM (2003) Increased resistance to pod shatter is associated with changes in the vascular structure in pods of a resynthesized *Brassica napus* line. J Exp Bot 54:1919–1930
5. Kuai J, Sun Y, Liu T, Zhang P, Zhou M, Wu J, Zhou G (2016) Physiological mechanisms behind differences in pod shattering resistance in rapeseed (*Brassica napus* L.) varieties. PLoS ONE 11:e0157341. https://doi.org/10.1371/journal.pone.0157341
6. Child RD, Chauvaux N, John K, Ulvskov P, Van Onckelen HA (1998) Ethylene biosynthesis in oilseed rape pods in relation to pod shatter. J Exp Bot 49:829–838
7. Agrawal AP, Basarkar PW, Salimath PM, Patil SA (2002) Role of cell wall-degrading enzymes in pod-shattering process of soybean, *Glycine max* (L.) Merrill. Curr Sci 82:58–60
8. Gan Y, Malhi SS, Brandt SA, McDonald CL (2008) Assessment of seed shattering resistance and yield loss in five oilseed crops. Can J Plant Sci 88:267–270
9. Kadkol GP, Beilharz VC, Halloran GM, MacMillan RH (1986) Anatomical basis of shatter resistance in the Oilseed *Brassicas*. Aust J Bot 34:595–601
10. Spence J, Vercher Y, Gates P, Harris N (1996) Pod shatter in *Arabidopsis thaliana, Brassica napus* and *B. juncea*. J Microsc 181:195–203
11. Liljegren SJ, Leslie ME, Darnielle L, Lewis MW, Taylor SM, Luo R, Geldner N, Chory J, Randazzo PA, Yanofsky MF, Ecker JR (2009) Regulation of membrane trafficking and organ separation by the NEVERSHED ARF-GAP protein. Development 136:1909–1918
12. Liljegren SJ, Ditta GS, Eshed Y, Savidge B, Bowman JL, Yanofsky MF (2000) Shatterproof *MADS*-box genes control seed dispersal in *Arabidopsis*. Nature 404:766–770
13. Ostergaard L, Kempin SA, Bies D, Klee HJ, Yanofsky MF (2006) Pod shatter- resistant *Brassica* fruit produced by ectopic expression of the *FRUITFULL* gene. Plant Biotechnol J 4:45–51
14. Rajani S, Sundaresan V (2001) The *Arabidopsis* myc/bHLH gene *ALCATRAZ* enables cell separation in fruit dehiscence. Curr Biol 11:1914–1922
15. Sorefan K, Girin T, Liljegren SJ, Ljung K, Robles P, Galvan-Ampudia CS, Offringa R, Friml J, Yanofsky MF, Ostergaard L (2009) A regulated auxin minimum is required for seed dispersal in *Arabidopsis*. Nature 459:583–586
16. Girin T, Stephenson P, Goldsack CMP, Kempin SA, Perez A, Pires N, Sparrow PA, Wood TA, Yanofsky MF, Ostergaard L (2010) *Brassicaceae INDEHISCENT* genes specify valve margin cell fate and repress replum formation. Plant J 63:329–338
17. Ogawa M, Kay P, Wilson S, Swain SM (2009) *Arabidopsis* dehiscence zone *polygalacturonases* (*ADPG1*), *ADPG2* and *QUARTET2* are polygalacturonases required for cell separation during reproductive development in *Arabidopsis*. Plant Cell 21:216–233
18. Mitsuda N, Iwase A, Yamamoto H, Yoshida M, Seki M, Shinozaki K, Ohme-Takagi M (2007) NAC transcription factors, *NST1* and *NST3*, are key regulators of the formation of secondary walls in woody tissues of *Arabidopsis*. Plant Cell 19:270–280
19. Liljegren SJ, Roeder AH, Kempin SA, Gremski K, Ostergaard L, Guimil S, Reyes DK, Yanofsky MF (2004) Control of fruit patterning in *Arabidopsis* by *INDEHISCENT*. Cell 116:843–853
20. Ferrandiz C, Liljegren SJ, Yanofsky MF (2000) Negative regulation of the *SHATTERPROOF* genes by *FRUITFULL* during *Arabidopsis* fruit development. Science 289:436–438
21. Alonso-Cantabrana H, Ripoll JJ, Ochando I, Vera A, Ferrándiz C, Martínez-Laborda A (2007) Common regulatory networks in leaf and fruit patterning revealed by mutations in the *Arabidopsis ASYMMETRIC LEAVES1* gene. Development 134:2663–2671
22. Hay A, Barkoulas M, Tsiantis M (2006) *ASYMMETRIC LEAVES1* and auxin activities converge to repress *BREVIPEDICELLUS* expression and promote leaf development in *Arabidopsis*. Development 133:3955–3961

23. Jenkins ES, Paul W, Craze M, Whitelaw CA, Weigand A (1999) Dehiscence-related expression of an *Arabidopsis thaliana* gene encoding a polygalacturonase in transgenic plants of *Brassica napus*. Plant Cell Environ 22:159–167
24. Raman H, Raman R, Kilian A, Detering F, Carling J, Coombes N, Diffey S, Kadkol G, Edwards D, McCully M, Kumar P, Parkin IAP, Batley J, Luckett DJ, Wratten N (2014) Genome-wide delineation of natural variation for siliqua shatter resistance in *Brassica napus*. PLoS ONE 9:e101673
25. Raman R, Qiu Y, Coombes N, Song J, Kilian A, Raman H (2017) Molecular diversity analysis and genetic mapping of pod shatter resistance loci in *Brassica carinata* L. Front Plant Sci 8:1765
26. Sra SK, Sharma M, Kaur G, Sharma S, Akhatar J, Sharma A, Banga SS (2019) Evolutionary aspects of direct or indirect selection for seed size and seed metabolites in *Brassica juncea* and diploid progenitor species. Mol Biol Rep 46:1227–1238
27. Kaur P, Banga S, Kumar N, Gupta S, Akhtar J, Banga SS (2014) Polyphyletic origin of *Brassica juncea* with *B. rapa* and *B. nigra* (*Brassicaceae*) participating as cytoplasm donor parents in independent hybridization events. Am J Bot 101:1157–1166
28. Liu XY, Macmillan RH, Burrow RP, Kadkol GP, Halloran GM (1994) Pendulum test for evaluation of rupture strength of seed pods. J Tex Stud 25:179–189
29. Kadkol GP (2009) *Brassica* shatter-resistance research update. In: Proceedings of 16th Australian research assembly on Brassicas conference, Ballarat Victoria, Australia, pp 104–09
30. Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. Focus 12:13–15
31. Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS ONE 6(5):e19379. https://doi.org/10.1371/journal.pone.0019379
32. Duitama J, Quintero JC, Cruz DF, Quintero C, Hubmann G, Foulquie-Moreno MR, Verstrepen KJ, Thevelein JM, Tohme J (2014) An integrated framework for discovery and genotyping of genomic variants from high-throughput sequencing experiments. Nucleic Acids Res 42(6):e44. https://doi.org/10.1093/nar/gkt1381
33. Roshyara NR, Scholz M (2014) fcGENE: A versatile tool for processing and transforming SNP datasets. PLoS ONE 9:e97589. https://doi.org/10.1371/journal.pone.0097589
34. Browning BL, Browning SR (2007) Efficient multilocus association testing for whole genome association studies using localized haplotype clustering. Genet Epidemiol 31(5):365–375
35. https://github.com/XiaoleiLiuBio/MVP
36. Lipka AE, Tian F, Wang Q, Peiffer J, Li M, Bradbury PJ, Gore MA, Buckler ES, Zhang Z (2012) GAPIT: genome association and prediction integrated tool. Bioinformatics 28:2397–2399
37. Götz S, García-Gómez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, Robles M, Talón M, Dopazo J, Conesa A (2008) High-throughput functional annotation and data mining with the Blast2go suite. Nucleic Acids Res 36:3420–3435
38. Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. Mol Ecol 14:2611–2620
39. Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES (2007) TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23:2633–2635
40. Goudet J (2001) FSTAT, a program to estimate and test gene diversities and fixation indices (ver. 2.9.3). https://www2.unil.ch/popgen/softwares/fstat.htm
41. Excoffier L, Hofer T (2009) Detecting loci under selection in a hierarchically structured population. Heredity 103:285–298
42. Beaumont MA, Nichols RA (1996) Evaluating loci for use in the genetic analysis of population structure. Proc R Soc Lond B 263:1619–1626
43. Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Res 10:564–567
44. Liu J, Wang J, Wang H, Wang W, Zhou R, Mei D, Cheng H, Yang J, Raman H, Hu Q (2016) Multigenic control of pod shattering resistance in Chinese rapeseed germplasm revealed by genome-wide association and linkage analyses. Front Plant Sci 7:1058
45. Wen Y, Fu T, Tu J, Ma C, Shen J, Zhang S (2008) Screening and analysis of resistance to silique shattering in rape (*Brassica napus* L.). Acta Agron Sin 34:163–166
46. Zhong R, Lee C, Ye ZH (2010) Evolutionary conservation of the transcriptional network regulating secondary cell wall biosynthesis. Trends Plant Sci 15:625–632
47. Dinneny JR, Weigel D, Yanofsky MF (2005) A genetic framework for fruit patterning in *Arabidopsis thaliana*. Development 132:4687–4696
48. Faris JD, Fellers JP, Brooks SA, Gill BS (2003) A bacterial artificial chromosome contig spanning the major domestication locus Q in wheat and identification of a candidate gene. Genetics 164:311–321
49. Gu Q, Ferrandiz C, Yanofsky MF, Martienssen R (1998) The *FRUITFULL* MADS box gene mediates cell differentiation during *Arabidopsis* fruit development. Development 125:1509–1517
50. Mandel MA, Yanofsky MF (1995) The *Arabidopsis AGL8* MADS-box gene is expressed in inflorescence meristems and is negatively regulated by *APETALA1*. Plant Cell 7:1763–1771
51. Lin RC, Park HJ, Wang HY (2008) Role of *Arabidopsis* RAP2. 4 in regulating light-and ethylene-mediated developmental processes and drought stress tolerance. Mol Plant 1:42–57
52. Meakin PJ, Roberts JA (1990) Dehiscence of fruit in oilseed rape (*Brassica napus* L.) II. The role of cell wall degrading enzymes and ethylene. J Exp Bot 41:1003–1011
53. Zhou Y, Lu DF, Li CY, Luo JH, Zhu BF, Zhu JJ, Shangguan YY, Wang ZX, Sang T, Zhou B, Han B (2012) Genetic control of seed shattering in rice by the *APETALA2* transcription factor *SHATTERING ABORTION1*. Plant Cell 24:1034–1048
54. He H, Bai M, Tong P, Hu Y, Yang M, Wu H (2018) *CELLULASE6* and *MANNANASE7* affect cell differentiation and silique dehiscence. Plant Physiol 176:2186–2201
55. Ehlers K, Bhide AS, Tekleyohans DG, Wittkop B, Snowdon RJ, Becker A (2016) The *MADS* box genes ABS, *SHP1*, and *SHP2* are essential for the coordination of cell divisions in ovule and seed coat development and for endosperm formation in *Arabidopsis thaliana*. PLoS ONE 11:e0165075
56. Zhang K, Akey JM, Wang N, Xiong M, Chakraborty R, Jin L (2003) Randomly distributed crossovers may generate block-like patterns of linkage disequilibrium: an act of genetic drift. Hum Genet 113:51–59